

INFO SUPPORT STAFF

Continuing

Intelligence

Methods

*file copy
D-116 file*

[REDACTED]

25X1A

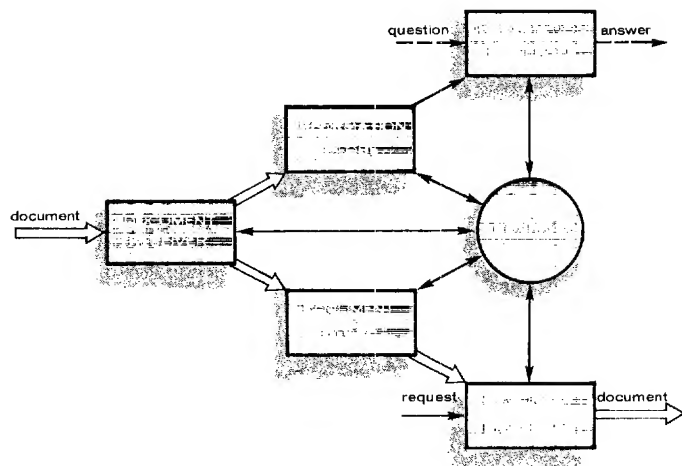
[REDACTED]

25X1C

INTELLIGENCE METHODS CONFERENCE

[REDACTED]

25X1C



SENDER WILL CHECK CLASSIFICATION TOP AND BOTTOM			
UNCLASSIFIED	CONFIDENTIAL	SECRET	
OFFICIAL ROUTING SLIP			
TO	NAME AND ADDRESS	DATE	INITIALS
1	IHC Support Staff		
2	Room 4E24, Headquarters		
3			
4			
5			
6			
	ACTION	DIRECT REPLY	PREPARE REPLY
	APPROVAL	DISPATCH	RECOMMENDATION
	COMMENT	FILE	RETURN
	CONCURRENCE	INFORMATION	SIGNATURE
Remarks: <p style="text-align: center;">As requested a copy for your retention.</p>			
FOLD HERE TO RETURN TO SENDER			
FROM: NAME, ADDRESS AND PHONE NO. 2927			DATE
			19Aug69
UNCLASSIFIED	CONFIDENTIAL	SECRET	

FORM NO. 1-67 **237**

Use previous editions

GPO : 1968 O - 297-542


(40)

STATSPEC

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

CONTROLLING INTELLIGENCE INFORMATION


Director of Intelligence Support
Directorate of Intelligence
Central Intelligence Agency

25X1A

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

SECRET

Acknowledgment

25X1A

A number of individuals have in one way or another contributed to this paper, including [REDACTED]

25X1A

[REDACTED] (OCS) with particular reference to Appendix B, and Project CHIVE personnel with particular reference to Appendix C. [REDACTED]

25X1A

(OCR) deserves special mention for major support in the preparation of the report as a whole, and my secretary, [REDACTED] for the preparation of the manuscript.

25X1A

P.A.B.

SECRET

There are some things which cannot
be learned quickly, and time, which is all we have,
must be paid heavily for their acquiring.

ERNEST HEMINGWAY

SECRET

TABLE OF CONTENTS


INTRODUCTION	1
I. TRENDS	2
II. IMPACT	8
III. CONTROLS	12
A. Management Controls	12
B. Selectivity	13
C. Reduction	15
D. Division of Labor	18
E. Optimizing the System	20
IV. AUTOMATION	24
V. CHALLENGE OF ASSUMPTIONS	30
APPENDICES: A. References	34
B. Selected Computer Applications and Techniques	38
C. CIA's Project CHIVE	44

SECRET

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

ILLUSTRATIONS & TABLES

. Percentage Increases in Unique Item Intelligence Documents	2a
. Growth in Receipts of Individual Title Open Literature Items	3
. Page Production of Translated Foreign Intelligence Materials	4
. Current Intelligence Production	6a
. 	12a
. Schematic of Community On-line Intelligence System (COINS)	21a
. Schematic of Computer-based Library of the Future	25a
. Relation of Amounts of Literature to Capacities of Computer Memories	25b
. Illustrative Model of the Intelligence Process:	
- Collection	27a
- Processing, Analysis, Synthesis	27b
. Project CHIVE	
- Current System	47a
- CHIVE System	47b
- CHIVE Functional Overview	47c
- CHIVE Information Language	47d

STATSPEC

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

~~SECRET~~

CONTROLLING INTELLIGENCE INFORMATION

INTRODUCTION

Beginning with the first Methods Conference, we have considered some aspect or other of the intelligence information problem. This is not surprising inasmuch as information is our business.

At [REDACTED] eight years ago, [REDACTED] provided us with useful thoughts on guiding collection.^{1/} A colleague at that time referred to [REDACTED] consideration of this particular aspect as contemplating "the tortured progress of a complex organism in getting its food from hand to mouth."

25X1A

25X1A

At that same conference, we also considered the processing of information, and ventured some judgments about the nature and direction of prospective changes, including the impact of automation.^{2/}

At Washington, four years ago, [REDACTED] brought us up to date on the state of the art of processing information automatically.^{3/} And I injected a corrective to the overly optimistic forecast made at the previous conference concerning the application in the specialized field of machine translation.^{4/}

25X1A

By and large these earlier reports and findings have held up well and remain valid. There would therefore seem to be little profit in an attempt merely to update what has previously been said. What would seem worthwhile is to assess the impact on intelligence of the increasing volume of information available. We should also strive to formulate guidelines which might be useful to our superiors for coping with this so-called information explosion problem.

~~SECRET~~

In the space of one hundred and seventy-six years the Lower Mississippi has shortened itself two hundred and forty-two miles. That is an average of a trifle over one mile and a third per year. Therefore, any calm person, who is not blind or idiotic, can see that in the...Silurian Period... the Lower Mississippi River was upwards of one million three hundred thousand miles long, and stuck out over the Gulf of Mexico like a fishing-rod. And by the same token any person can see that seven hundred and forty-two years from now the Lower Mississippi will be only a mile and three quarters long, and Cairo and New Orleans will have joined their streets together.... There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact....

MARK TWAIN

I. TRENDS

Scholars have long cherished the right to complain about their inability to cope with the amount of information at their disposal. In 1613, it was Barnaby Rich who referred to the multiplicity of books as one of the diseases of his age. Books, he said, "doth so overcharge the world that it is not able to digest the abundance of idle matter that is every day hatched and brought forth...."^{5/} In 1945, it was Vannevar Bush who expressed concern over the startling growth in scientific literature and described the investigator as "staggered by the findings and conclusions of thousands of other workers--conclusions which he cannot find time to grasp, much less remember, as they appear."^{6/}

Since then, whether in or out of Government, in scholarly journals or during budget reviews, it is a favorite theme of analysts to lament the unmanageability of the in-basket, prompted, it is said, by the information explosion. Most are quite ready to testify that this situation is getting worse despite improvement in methods for controlling information.

The reason for this is not hard to find. We recently polled our collection, analysis, and information handling staffs to determine, in gross terms, what the current trends in intelligence information appeared to be and where the effects of change in data flows were most severely felt. We indeed found volume trends for all types of information of interest to intelligence invariably up. This is so for the large amount of open source literature, that is, information published worldwide in newspapers, periodicals, and books. It is so for the amount of government field reports and the finished analyses produced by the intelligence community itself. It is very much the case as to technical collection (SIGINT and overhead photography).

As can be seen, where we deal with receipts of hard copy materials alone, the percentage increase since 1950 is very considerable. Only in the case of collateral intelligence documents has the curve flattened since 1960 (see following chart).

25X1A

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

~~SECRET~~

One of the most startling increases in information receipts has been in the open literature of books, serials, and newspapers. Volume figures rose sharply in the years immediately following the death of Stalin and reflect the relaxation of Soviet control over external subscriptions and the early receipt of Chinese S&T literature. Subsequent increases indicate the continued upward movement of S&T literature and the enlarged volumes in free world publications and in publications from the emerging areas of Africa, Latin America, and Asia which must be screened to identify information of intelligence interest.

Growth in Receipts of
Individual Title Open Literature Items

	<u>1950</u>	<u>1960</u>	<u>1966</u>	<u>1970</u>
English S&T	2,100	4,700	5,300	6,000
English Soc- iological	5,300	10,100	11,900	13,700
Foreign S&T	1,900	8,500	9,600	10,800
Foreign Soc- iological	6,700	16,200	19,100	22,000

Understandably, translation of open source foreign language materials has followed a similar trend, showing a steady rise beginning with the relaxation of Soviet control over subscriptions, and continuing with the rise in interest in published materials from the emerging areas. Changing requirements have brought entire new areas of foreign language materials under scrutiny, including those of Africa, Latin America and Cuba, and South and East Asia. As an indication of the shift in emphasis, our Foreign Documents Division, which in 1960 had only two translators assigned to Spanish-language exploitation, now has 14 people working on Latin American and Cuban press materials and is issuing a daily Latin American press exploitation report.

Our total annual page production in translated materials has doubled in the past six years, and it is anticipated that by 1970 this output could reach a half million pages a year.

~~SECRET~~

SECRET

Page Production of
Translated Foreign Language Materials

<u>1950</u>	<u>1960</u>	<u>1966</u>	<u>Estimated</u> <u>1970</u>
59,000	212,000	425,000	500,000

To radio monitors, the information explosion is no myth. A few examples illustrate the dramatic proliferation in broadcasting.

In 1956, Radio Moscow broadcast about 115 hours a day to foreign audiences; this year the figure is 213 hours. Ten years ago Moscow broadcast in 40 languages, today it broadcasts in 68. Then it carried about 77,000 "commentaries" or long items each year; now the annual figure is 177,000.

Radio Peking's output also has skyrocketed. In 1956, it broadcast 47 hours a day in 15 languages, today it broadcasts almost four times as much (166 hours) in twice as many languages (32). Ten years ago the Chinese broadcast 25,000 commentaries annually; that output has now nearly quadrupled (97,000).

This increase in volume has been accompanied by new demands for coverage as one-time Soviet "satellites," whose media could largely be dismissed as echoes of Moscow, have become independent forces to be reckoned with and as new problem areas in Latin America, Africa, and South Asia have focused fresh attention on radios in these areas.

25X1D

SECRET

SECRET

25X1D

In the field of reconnaissance, there is in fact an information explosion affecting the processor, producers, and consumers of photo intelligence. Among the various reasons for such an explosion, the most important perhaps, are the increased use of our ever-improving reconnaissance capabilities, an increased

25X1B

general widening of interests into areas previously neglected or ignored by intelligence.

25X1B

25X1D

Prior to 1956 the primary source of this data was captured World War II photography. With the advent of means to collect aerial photographs of those areas of vital interest to us the situation quickly shifted from one of lack of photography to one of properly exploiting, on a timely basis, the wealth of photography available. Volumes currently processed are many times the 1960 volumes. For example, during fiscal year 1965 an excess

25X1D

25X1B

SECRET

SECRET

25X1B



One factor not accounted for here is the rise in copy distribution. Our documents staffs are disseminating to more points than ever before. About 35,000 classified intelligence documents flow into our dissemination center each day (in about 12 copies each) and are disseminated to more than 200 offices or individuals. In the case of open literature the copy volume in 1966 exceeded 1.6 million (1,237,000 daily newspapers; 270,600 journals; 114,500 books). This is a three-fold increase over 1950. By 1970 the figure will be two million.

In addition, figures indicating receipt or production of unique items do not reveal the increase in number of pages issued yearly. There has been a steep rise in page production for certain series of reports. Illustrating this trend is the page production of translated materials cited earlier and the output of "impressions" (individual page reproductions) from our current intelligence office. This latter figure rose steeply from an annual output of seven million in 1955 to more than 30 million on 1965. While much of this latter increase represents production for enlarged readership rather than contributions to the information pool, a good part represents additions in the form of analytical commentary (see following chart).

One curious development, not usually taken into account, is that the computer itself is currently aggravating rather than alleviating the information problem. The brute-force power of EDP has in many cases been substituted for adequate analysis of a data processing problem. This has been the case for many new computer applications. It is only after the second or third expensive iteration that a sensible perspective is obtained.

SECRET

25X1A

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

We see instances where basic input data is replicated, reformatted, resorted, to the point where the output volume exceeds input by perhaps 100 times. It seems evident that the same methods applied to harnessing the information explosion elsewhere in the intelligence process must also be applied--perhaps more carefully--in the computer environment.

All in all a significant part of the information problem is of our own creation. The overall magnitude for the American intelligence community of the information handling problem was given some clarification in a survey of information inventories and flows. This study estimated that we annually produced or handled 20 thousand individual series, in ten million issues, published in 150 million copies. While the amount originated by intelligence agencies is only a fraction of these totals, it is still considerable. For example, one component of one agency alone produces some 44,000 information reports annually. An average of 90 copies of each is distributed initially for a total of four million copies.

Add to the above the fact that the advent of efficient office copying machines has encouraged secondary reproduction by recipient offices to absurd proportions.

It is little wonder that intelligence analysts have come to feel that they have in the space of a few years moved from a condition of having too little to having too much information.

Where is the Wisdom we have lost
in knowledge?

Where is the knowledge we have
lost in information?

T. S. ELIOT

SECRET

II. IMPACT

At this point it is well to make some distinctions. Not all analysts are equally effected by changes in information availability and receipts. Moreover, the demands made by policy makers and planners upon production analysts are themselves subject to considerable change. The relationship between the causes of change in the demand for intelligence and in the availability of information upon which that intelligence is based may be a significant one. But this has not been the matter of serious study and little about it is known.

It is important therefore to resist the tendency to view the aggregate receipts from all sources--teletype, radio and TV, ground and overhead photography, COMINT, maps, books, newspapers, and serials--as equally affecting all elements of the intelligence process. Analysts throughout the intelligence cycle are affected by the changes in the information flow according to their area and type of assignment. It is not often the case that an individual analyst requires all types of incoming materials to do his job. One of our economic research areas, for example, indicates that open source and recee information comprise the bulk of the significant data on which their analysts rely. Specialization by area and topic also tends to reduce to manageable proportions the volume of sources a analyst must examine. In addition, the production analyst is shielded from the full force of the information flood by an elaborate filtering and requirements process, and by a wide variety of indexing, abstracting, collating, and reference services.

Information deriving from new collection means or from open literature has had its greatest impact on the military, economic, and scientific/technical areas with the main emphasis falling in the Soviet Bloc. However, with the exception of open source information, the bulk of the new material arrives as raw data of which only a fraction expresses itself in the form of published materials through which the production analyst must make his way. It is probably a safe generalization to say that, with certain exceptions, many research or production analysts are not acutely aware of living in the midst of an

SECRET

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

information crisis. Indeed in some cases we may find an information drought in an area of plenty, largely for organizational or logistical reasons. For example, the main reason for liberal copy creation and dissemination is to get the word to those who need to know. This has to be done through set channels of communications. The resources devoted to the operation of these channels must not be disproportionately high to the resources allocated to the substantive use made of the materials transmitted. A point is therefore soon reached beyond which the system will not carry additional traffic. In this squeeze dissemination policy is reviewed, "marginal" accounts are dropped, and with the passage of time a given analyst will not be receiving what he needs. What's more he won't know that he isn't.

This indeed emerged as the main conclusion of a study undertaken to find ways and means to staunch the flow of paper to the analyst. We thus have the paradox of too much paper for the analyst to handle but not all that he ought to see to be on top of his job.

The scientific analyst is one who is all too well aware of the difficulty of staying on top of the S&T literature. Where, formerly, he screened the literature himself, he now relies on translation services to scan, select and abstract material for him. By 1970 the volume of S&T material will be overwhelming, and mechanical means will be required to scan and select important items and collate information.

The full force of the flood is felt by analysts working in collection, data processing, or in reference activities. It is from these quarters that the greatest anxiety about an "information explosion" has been voiced, for it is these processors (intelligence support analysts) who face the huge data reduction problem or must cope with the cumulative effects of high volumes of information on their ability to transmit, to disseminate, to store, and to retrieve.

Certain changes in the climate in which analysts work tend to create a further effect of large-scale expansion of the information load. Among these changes are the emergence of new topics or world areas of intelligence

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

significance. Fifteen years ago the world's sovereign states numbered 80. Today there are 128. A rise in topics of interest has accompanied the growth of nations; a decade ago we were not concerned with a French or Egyptian missile program, [REDACTED]

[REDACTED] A gazetteer on Antarctica is about to be published with over 11,000 place names. Hundreds of new personalities now play important political roles.

There is a tendency to spread analysts fairly thin in assigning work on new areas. In addition, new resources, and some resources diverted from established activities, have been required to man new collection, requirements, and data handling activities. Technical innovations, particularly in communications, have brought the reporting and response to events into a "real time" mode. Events are reported as they are happening, and the report from the field not only informs but often generates requests, creating an information-request-information spiral.

In general the production analyst has made few changes in his traditional approach to his tasks. In areas burdened by manpower shortages or by heavy document or ad hoc production loads he may tend to read more casually, to treat more information as background, to file fewer documents, and cancel subscriptions. One office has placed its filing activities on three levels of activity: active research files, files which are maintained but on which no research time is expended, and, for areas not being actively followed, files which include only the finished reports of others. In some instances analysts will cancel receipt of material on inactive fields but will expect the information to be collected and filed by central reference against the time when the field again becomes active or when basic studies must be updated.

While there is a tendency to place more reliance on central reference functions, particularly on such specialized services as those provided by biographic and installation files, there is apprehension about the lack of an all-source capability and about the inability to retrieve materials by subject and area. In most cases, the analyst continues to regard his active file as the best, relying on central reference service for backup,

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

for developing information on new areas and topics, or as a check on dissemination. Reliance on summaries and handbooks produced by central reference or other research areas has increased, and these products fill a need for basic reference materials.

While analysts have been active in changing requirements and in making more strict definitions of their missions, they are often called upon on short notice to answer questions on matters which they have not been able to follow in depth. Nearly all production offices complain of the disruptive and time-consuming nature of the current requirement to divert time from production or analysis into ad hoc activities. One component has indicated that some analysts are concerned with work on ad hoc problems to the point of falling out of touch with their research fields. This situation indicates a need to review mission statements and organizational philosophy in the light of current response requirements.

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

...it takes all the running you can
do to stay in the same place.
If you want to get somewhere
else you must run at least
twice as fast as that!

THROUGH THE LOOKING GLASS

SECRET

III. CONTROLS

Thus far we have seen that an information problem exists, and that its impact upon analysts varies, but is considerable, especially upon analysts in the support services. Let us now look at some of the techniques which we actually employ or are devising to control information.

A. Management Controls

A number of basic and direct techniques are always at hand to control the data flow. These are in the area of exclusive or arbitrary controls, e.g., cut manpower/increase manpower; cut budgets/increase budgets; apply output limitations; reorganize data processing activities; eliminate editorial controls. All of these have been applied at one point or another in the intelligence cycle. One successful application of such an arbitrary control is

STATSPEC

Reduction of manpower and budgets is an often used control. It, also, imposes high selectivity standards and has the advantage of requiring the victim to look sharply at his mission description, not solely from the point of view of what his mission is but also is not.

The expansion of the all-source concept has provided gains and still shows considerable promise. The increase in cleared personnel has allowed a free exchange of views and documents, has made it easier to obtain collated information, has bridged the gaps caused by compartmented, limited source activities. It offers an overview of what is known on a subject from open source materials all the way through the products of [REDACTED] systems. It has been of particular value in assisting rapid evaluation of recce data.

25X1D

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

STATSPEC

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

Reorganization is a way of life in the intelligence community. All too often, however, reorganization is inspired not by an examination of existing responsibilities and work patterns, but by the emergence of new requirements and priority areas. At present our production directorate is examining a number of activities to seek more effective combinations of functions and areas of assignment of responsibility.

B. Selectivity

One major organizational advance was the formation in 1963 of a Collection Guidance Staff (CGS) to serve as a focus for coordination of collection activities, particularly those involving high investment, technical methods. One feature of the CGS is an all-source requirements registry which can match information to intelligence needs, recall reporting on a subject or area for evaluation by analysts, and identify reporting which is not related to stated intelligence needs. The coordination of requirements and output from various collection means has enabled us to limit collection in many areas. In addition, the gain by the user in technical knowledge of collection capabilities and limitations has resulted in more refined and meaningful guidance to the collector.

Advances in communications technology, teletype and long-distance facsimile, have not only made information more timely but have enabled the analyst to get to the collector much faster with much more information than ever before. This improved dialog has made collectors more responsive to guidance and has made analysts more appreciative of good collection guidance. There has also been increased initiative on the part of the collector in reviewing requirements and reporting needs.

In photo collection, attempts to exclude marginal information at the source have centered on trying to hold down the number of reconnaissance targets. These efforts, however, are complicated by the desire of each area or subject specialist to regard his particular set of targets as valid, necessary, important, or critical. In addition, there is a reluctance by all consumers of photo intelligence to permit even marginal, poor quality photography to go unexamined and uninterpreted. The main technique currently being used to place both the collection and

~~SECRET~~

exploitation (interpretation) of reconnaissance photography in order of importance is to accept the requirements of an inter-agency committee or working group (where available) which is concerned with a particular problem rather than to accept those of individual analysts. While effective in providing priorities, this has had little effect on the total volume of material collected in this relatively new area, though it has been more successful in dealing with other sources.

STATSPEC

One of our offices has made good progress in controlling dissemination of hard copy materials by encouraging its analysts to accept search responses in microform. Aperture cards, containing up to eight microimage document pages, are sent to requesters who must return the cards when no longer needed. This system has had the advantages of economy and security, of reducing the number of hard copy documents being filed, of reducing actual file sizes, and of good customer feedback--by punching out a hole in the aperture card the user may request hard copy or indicate that the document ought to be purged from the file. Other forms of microform service are under examination, including the possibility of reducing reference and indexer aids to rapid-lookup microform systems.

~~SECRET~~

SECRET

C. Reduction

1. Title and accessions lists

Various types of non-mechanical data reduction techniques are being used with good results. One of the simplest of these is the use of title listings or announcement bulletins to inform users of the availability of materials, avoiding the need for scanning, abstracting, translating, or dissemination of large volumes of material. We have a system through which the tables of contents of selected domestic and free world scientific journals are copied and sent to users in accordance with their established coding requirements. The Russian Book List gives translated titles of incoming Soviet books, informing the analyst of their availability and enabling him to determine which ones he wishes to review or have translated. The Consolidated Translation Survey is a monthly record of translated items or items in the process of translation. It serves as an announcement of availability of translation of items of interest and as an antiduplication device. Accessions lists of maps, photos, and other limited copy or hard-to-disseminate items are in wide use.

2. Indexes

The application of a large array of indexing schemes is used to control large document collections and to reduce reading and search time. Machine-controlled indexes guide the user to information in COMINT and collateral documents and provide special controls over biographic and installation hard fact information.

In the area of published indexes, the Intelligence Publications Index provides the user with a monthly catalogue of finished intelligence publications and reports. Published indexes to selected foreign newspapers or radio broadcasts serve not only as guides to further exploitation but as press emphasis indexes and as a reference to important speeches or articles.

STATSPEC
STATSPEC

SECRET

SECRET

STATSPEC



Simple indexes of "keywords" which can be searched in combination are useful guides to large data collections and are also of value for dissemination. Keyword title listings are used to provide minimal control over the large volume of COMINT data and are valuable in freeing the data from the rigidity and ambiguity of a classification schedule. While it is unlikely that users will accept indexes as a substitute for dissemination in fields they are actively following, indexes may be accepted in place of documents on lesser priority topics, as a check on dissemination, or as a desk reference.

One such index, which will be issued in the near future, is the SKAN (Subject Keywork Announcement) index of incoming ChiCom documents. This keyword-out-of-context listing will be the first product of the computer system being implemented in our central reference facility. This index can be sorted by topic and area and special issues, of interest to individual analysts, prepared, and distributed.

3. Synthesis

Certain types of data known to be of general utility, particularly those reported from many sources, can be effectively synthesized and reduced in the form of listings, indexes, and summaries. Often these are prepared for rapid publication and distribution. An example of the publication of selected data elements is

25X1B [redacted] publication of [redacted] division. Finished intelligence publications are the most significant product of synthesis, with the NIE reflecting the largest reduction ratio.

25X1B

SECRET

SECRET

4. Abstracting

There is an expressed need for the production of more information containing abstracts or prepared in the form of abstracts as an aid to rapid review and identification of data of importance. Producers of reports provide service to screening by including introductory abstracts. These often contain enough information in the way of conclusions to stand alone or to save reading the entire report to determine its interest or value for filing. Abstracts of foreign literature, particularly books, are also required. In the vast area of S&T literature an active contract program, coordinated between CIA and the Department of Defense, arranges for the preparation of abstract translations of items from Soviet Bloc journals and monographs. The results of this load sharing are published together in topical serials (Bio-Medical Science, Chemistry, Electronics) and are made available to the scientific world through the translation service of the U.S. Department of Commerce. Not only do these abstracts overcome the barrier of many languages and technical vocabularies, obviate the need to maintain large numbers of individual subscriptions, but they enable the user to screen large masses of data rapidly and with ease.

In the central document files an abstract system known as DARE (from Document Abstract Reproducing Equipment) has been implemented. This system, by copying and reducing the first page of a document onto a card which can be sorted by EAM equipment, eliminates the need for tedious and costly human abstracting and often enables the user to determine at a glance whether or not the document meets his requirements without the need to work toward the document by way of an abstract statement. Since an estimated 31% of all information reports consist of a single page, while 34% are limited to two pages, savings in time and effort of both the processor and the user are significant.

5. Formatting

Preparation by the originator of data in a form suitable for immediate file input can be a major step in shortening reading and filing time. Good progress has been made on this front and a wide variety of programs

SECRET

SECRET

exist in which data is formatted in such a way as to be ready for filing upon receipt by the user. Our China Plant Data Report is an example of a collection format which structures significant data elements according to the logic of the user's file. Plant names are highlighted for rapid identification and screening, and the inclusion of only "hard fact" or meaning-bearing words (locations, production figures, floorspace) eliminates the need to translate and read much superfluous text.

There are a number of formatted sources containing biographic information which are received on contract or from other government agencies. Hard fact economic, technical, and geographical information is similarly prepared for [REDACTED] Recce collection, in addition to making requirements more specific, circumscribes the collector's response within rigid reporting formats.

Highly formatted data lends itself readily to computer processing. Key punching is simplified or flexotape can be prepared at the time of processing for later conversion to magnetic tape. A number of applications, which will be discussed later in this paper, are in process in this field.

D. Division of Labor

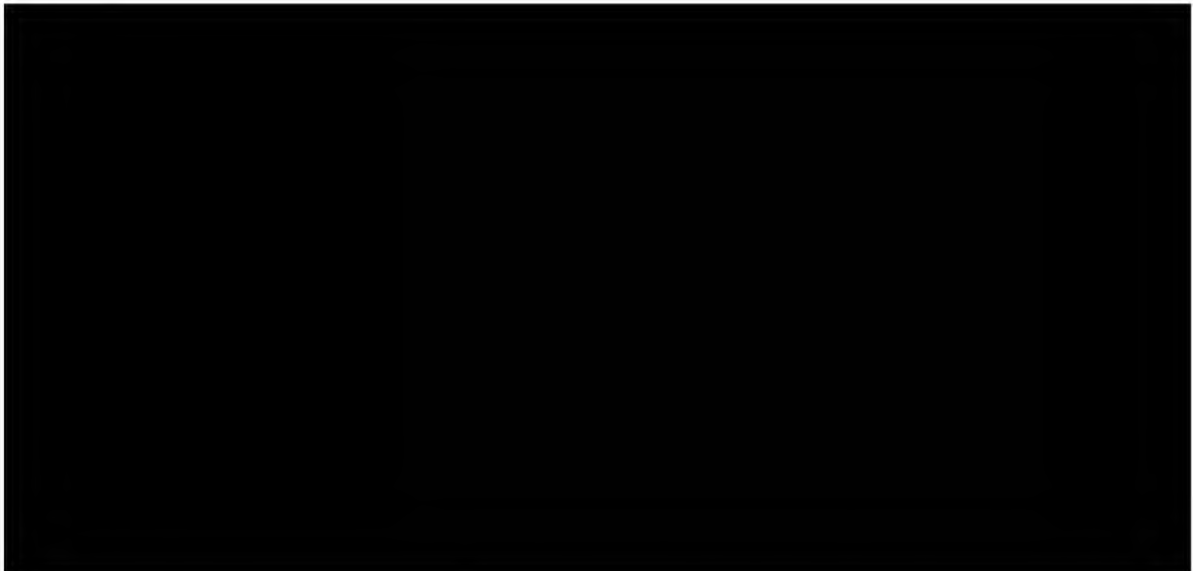
In a period of tight budgets and expanding requirements it is essential that duplication of effort be eliminated to the greatest extent possible, that cooperative ventures be developed whenever feasible, and that we rely on the work of others to accelerate the pace of our research and ease the burden of our personnel.

Obviously a clear delineation of mission responsibility and the elimination of overlap by organizational approaches offers promise in this area. The grouping of "like" functions into single organization entities is useful in providing more integrated production and less duplication of effort than is required by coordination processes. The bringing together of experts to comment on a given problem also offers opportunities for gains in time and efficiency. We have had some experience in this area with still another type of committee--the task force--which is often formed during a crisis or in anticipation of the

SECRET

SECRET

emergence of a critical problem. Such task forces have been in operation on Cuba, Vietnam, the Dominican Republic, China, and the Soviet succession issue. These bodies bring together experts from all intelligence disciplines--collection, analysis, reference--and provide a highly integrated, fully informed attack on the target problem. The degree of success which these operations have achieved probably has strong organizational implications and should be studied from that point of view. STATSPEC



There are other, more obvious forms of load sharing. Handbooks, statistical yearbooks, the finished work of Free World economists and the publications of research foundations and corporations provide reliable basic materials which eliminate the need for additional collection and analysis. A recently translated publication

25X1C



25X1B [redacted] dictionaries, gazetteers, and city directories, particularly in the Free World can be used and filed as received. Even Soviet [redacted] can be relied on to give at least the outlines of a [redacted] which could not easily be pieced together from the collection of many identifying items. Internally

25X1B
25X1B

SECRET

~~SECRET~~

25X1B produced handbooks--[REDACTED] geographic, general area survey--are widely used as basic aids to research, thus saving the time of many analysts in screening documents, maintaining files, and searching for information.

E. Optimizing the System

This catch-all can be said to cover a wide area of techniques and methods which could affect the environment within which information is handled, the relationships between different parts of a system, the form of the information itself, or man-machine relations. Most of the techniques are not unique in their application to information handling and I do not propose to dwell on them at length. They are of importance however. They can be used with expectation of realizing efficiencies in the activities studied.

1. Systems analysis

Techniques for optimizing a system would include systems analysis and operations research to study flow of data patterns, to locate bottlenecks, to identify overlapping functions, to eliminate unnecessary steps, to simplify channels of communications. Systems training for personnel is required to qualify them to bring about orderly change--and these should if at all possible be members of your own staff. Managers need such training also to enable them to make informed decisions about the applicability of automated techniques to their operations.

2. Standardization

The high volumes of today's information, and the rapid processing and response speeds required, indicate an increasing dependency on exchange of data in the form of tapes and graphics, the ability to rely on and remotely query the files of others, and the capability to collate data from many sources in a single product. This requires standardization, that is, speaking the same language and agreeing on the meanings of forms and terms. In addition, the requirements of computer processing are compelling us to seek common formats and languages to facilitate computer input and interface. Moreover, standardization permits a wider application of the concept of division of labor.

~~SECRET~~

~~SECRET~~

The difficulty of standardizing on a broad front is illustrated by the failure after 20 years to agree on a standard Cyrillic to Latin transliteration system. Standardization problems are among the issues on which USIB Committee on Documentation task teams have been meeting over the past year. Goals include standards for geographic and installation codes, for film storage media, and for report identification.

3. Program and organization planning

With program planning we hope to achieve a better job of allocating resources to information handling programs responsive to priority needs. By flexible organization we endeavor to anticipate requirements for organizational change.

4. Analyst communications

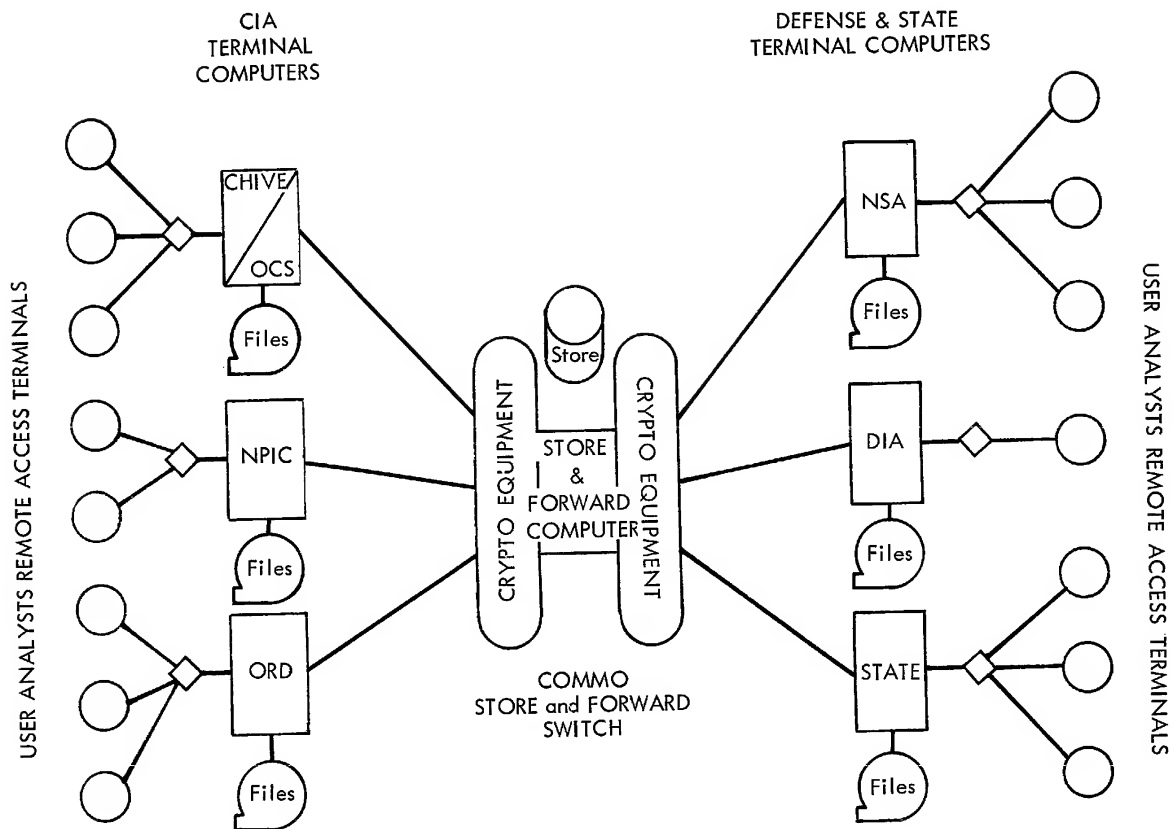
Analyst communications is another very important part of the system. Analyst files represent a refined research product ready for response on special problems. Analysts themselves represent a community of available experts competent to provide information or informed judgment. We need to improve the channels of communications between these experts and files, the means of locating what is known, who knows it, and how to bring the full range of available expertise to bear on a given problem.

In the realms both of sharing the burden and of analyst communications, U.S. intelligence has initiated an experimental Community On-Line Information System (COINS). This will permit an analyst in one agency, through a remote desk console, to query a file stored in the computer of another agency (see following schematic).

5. Upgrading the analyst

Upgrading the analyst is another promising means of handling information more successfully. The production analyst can be upgraded by more specialized training (including foreign language training which would reduce the load he places upon the translation facility), or by being given support (human or machine) designed to relieve him of the less productive aspects of his task. We are doing both.

~~SECRET~~



Schematic of
Community On-line Intelligence System
(COINS)

SECRET

We have long provided massive reference and support services to our production analyst. The present thrust is to make reference personnel a more active participant in the production process and the production analyst an active participant in a system design to support him.

In photo interpretation, for example, we have had good results in teaming the PI with the all-source analyst in work on specific problems. Our [REDACTED] support service has also worked well. The question is no longer whether there shall be both a political analyst and [REDACTED] in support, or a single intelligence officer to do both jobs. Rather the question is to determine the optimal relation between the two. Nor is the question whether the [REDACTED] shall or shall not receive machine support, but how that support is to be organized.

25X1B

25X1B

A recent user survey, conducted in the Department of Defense, provided statistical evidence for certain intuitive views we have held for some time. First, that specialists in information handling can not only fulfill user needs but often anticipate needs before the user is fully aware of the requirement. Secondly, in 68% of the DOD sampling, the user preferred to obtain a detailed analysis or a specific answer to his request rather than a set of abstracts to documents or, presumably, documents themselves. These survey results have meaning for our own activities in how we use our reference staffs; we need to gain a clearer view of the function and potential of the information specialist in creating new information through synthesis and correlation of available data, and of his responsibility for monitoring the user's process of information acquisition, use, and generation.

6. Providing user feedback

User feedback, whether related to collection, reference, or production, is an essential element of any system and one of the least satisfactory element of our own. Often analysts work for long periods completely in the dark as to how satisfactory their work is to users or to what purposes it is being put. There exist also problems of making the full range of services or functions of one office known to potential users in other office.

SECRET

Our reference staffs have been seriously concerned about this matter. Conversely, reference staffs have few ways of determining what the user's problems are except to the extent this information can be derived from request patterns or through the tedious and dubious process of surveying user needs. Often the capability of information personnel to tailor responses directly to user needs goes unused because of inadequate communications. It may be time to think about spending less time attending documentation meetings and more time studying modern retailing philosophy. The situation between the production analyst and his customer, the policy maker or planner, is also unsatisfactory, though we do have cases where unexceptionable relationships have been developed to the mutual benefit of producer and recipient.

SECRET

Cold blows the wind on my true love
And a few small drops of rain --

SECRET

IV. AUTOMATION

Quite recently the Saturday Review devoted a special issue to the potential of the automation revolution and its implications for our society.^{7/} This forecast is an impressive one. So impressed were the editors that they equated the computer with the atomic bomb as a technological development formidable enough to mark a turning point in human history.

A Newsweek report,^{8/} prophetically entitled "Good-by to Gutenberg," also gave readers a glimpse of things to come in information technology: a photosensitive crystal the size of a lump of sugar that is capable of containing images of 100,000 pages; a lensless photographic system which could lead to three-dimensional home television; a no-contact, no-pressure printing technique that can write on sand, print a message on a pizza, and put a trademark on a raw egg yolk.

Such concepts, and dozens of others are indeed incubating in industrial laboratories and being translated into hardware. This progress is being heralded as the final phase of the U.S.'s \$10 billion communications industry's transformation from the old Gutenberg ways. Marshall McLuhan for example predicts that books and newspapers will no longer exist; that publishing will become an active servicing of the human mind through research packages done to suit individual needs.^{9/}

To those of us who have read similar promises over the last 20 years, these words smack of extravagance. I do not doubt that amazing developments will continue to take place. I do doubt that we can count on them for early and revolutionary solutions to our information handling problems. The kind of problems to which I have referred earlier have a habit of staying well in front of innovations.

In CIA we are for example currently upgrading our computer facilities to third generation hardware infinitely superior to our initial gear. Yet the contribution of the computer to the task of producing intelligence is still both specialized and very limited. We have secure

SECRET

SECRET

telephone communications. But these are far from ideal, the number of instruments are few, the cost of each is high. Great strides have been made in our printing establishment. Still, the lapse between preparation of copy and its availability to the anxious reader can be measured in weeks rather than days for nonpriority items. Reproduction techniques have shown major gains. But material of intelligence interest received in such poor quality that it cannot be microfilmed runs in some categories as high as 20%. We have improved our training of people and our means of furnishing them with better instructions. This notwithstanding, 50% of the titles of some report series need to be rewritten in headquarters in order to reflect substantive content meaningfully.

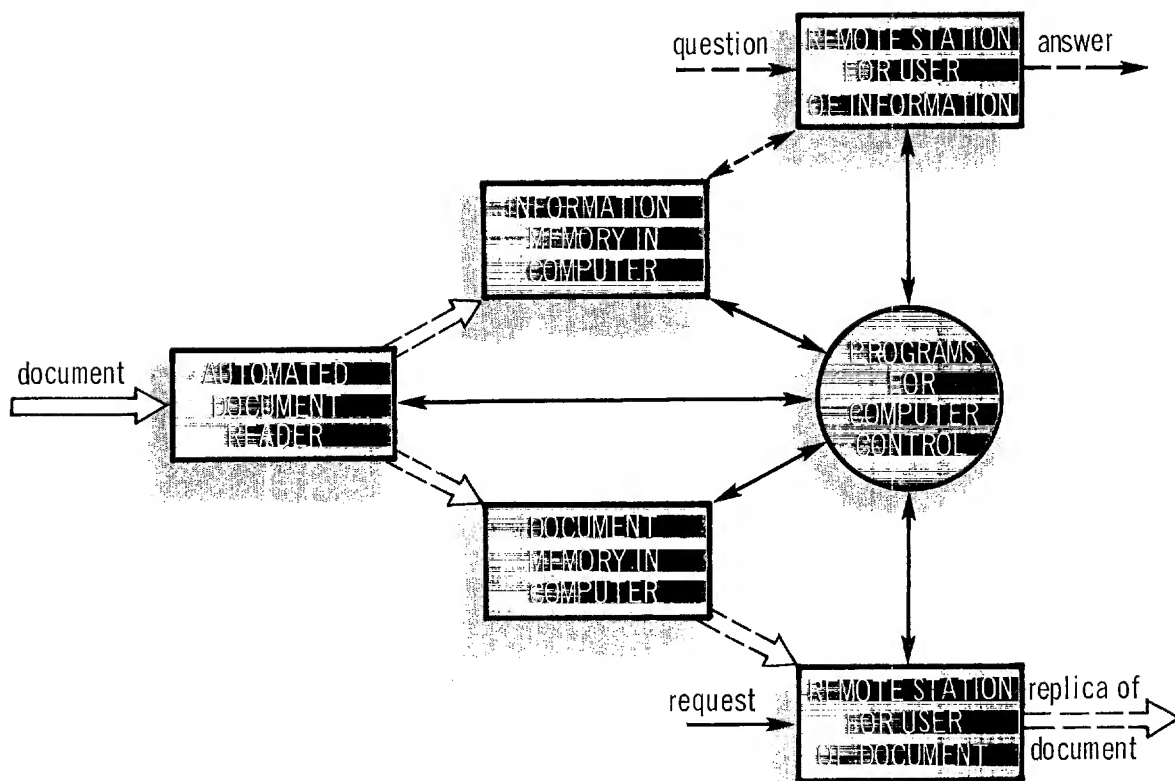
I do not cite these findings to ring the changes of gloom. I cite them as a call to caution against going overboard in diverting resources from the workable (if less than satisfactory) old to the untried new.

Having said that, I nevertheless urge that more attention be paid by intelligence to what is going on outside. As never before we have opportunities to capitalize on the work and ingenuity of others as a source of relief for some of our own problems. Much of the work done outside is solid and relevant. We ought therefore to pick-a-back whenever we can.

Let me mention a few of the more important outside activities I have in mind. Two are noteworthy in the academic community at large--INTREX^{10/} and EDUCOM. INTREX, which stands for "information transfer experiments," has been called progress toward a dial-a-thought world. INTREX is setting up an experimental laboratory to test ways of giving professors and students instant access to information. Xerography, film projection, and telephone communication between computer and user are planned. Basically experiments conducted will (1) attempt to automate and rationalize the functions of libraries, and (2) develop an information transfer network which is computer based (see following schematic).

EDUCOM,^{11/} the Interuniversity Communications Council, is an organization of over 30 universities in 20 states. It aims at evaluating the significance for higher education of electronic hardware (computers, light pens, graphic displays), and software (computer programs).

SECRET



SCHEMATIC OF COMPUTER-BASED LIBRARY OF THE FUTURE

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

**THIS TABLE RELATES AMOUNTS OF LITERATURE
TO CAPACITIES OF COMPUTER MEMORIES ***

World's total store of literature - Year 2000	10,000,000,000,000,000
World's total store of literature	1,000,000,000,000,000
All literature in science and technology	100,000,000,000,000
Capacity of fast random-access memory - Year 2000	10,000,000,000,000
Capacity of serial-access memory being developed	1,000,000,000,000
Literature in average field of science	100,000,000,000
Literature in average sub-field of science	
Capacity of fast random-access memory - Year 1980	10,000,000,000
Capacity of magnetic disk file	
Contents of Encyclopedia Britannica	1,000,000,000
	100,000,000
Information in average book	10,000,000
Capacity of fast random-access memory	
	1,000,000

*The scale measures the number of bits, or binary digits, of information that can be stored in the computer memory, or that are required to store the contents of the literature. All figures are estimates for 1964 except as noted otherwise. The capacity of fast random-access memory, now of the order of ten million bits, is shown in the last row.

~~SECRET~~

A number of individual university libraries have forward looking programs, Washington State^{12/} and Florida Atlantic^{13/} to name only two. The latter library has the distinction of being the first in the U.S. to introduce data processing methods and techniques into its operations from the very beginning. Washington State on the other hand is converting from traditional library methods to a totally on-line library system which offers multiple remote access to a single library record. Thus by sharing the time of the university computer (IBM 360/67), it will be possible to reduce typing substantially, eliminate duplicate manual files and have complete control of each item's location and status in the library system.

The value of these projects to us is that they are more in keeping with the size of projects which intelligence libraries may undertake. While much valuable data has been published about ways and means for automating the Library of Congress,^{14/} the sheer size of the holdings makes many of the parameters of that undertaking of little practical value to us.

The publications of professional engineers, documentalists and specialists in various aspects of the information handling industry are also increasingly more solid and relevant. Particularly useful books, for example, have been published by Licklider,^{15/} Bourne,^{16/} Becker and Hayes,^{17/} Greenberger,^{18/} Feigenbaum and Feldman,^{19/} and Borko.^{20/} The National Science Foundation's "Current Research and Development in Scientific Documentation"^{21/} serves to alert readers to on-going experimentation in information handling techniques. And the proceedings of professional meetings on this subject give first-rate coverage for those who cannot attend.^{22/}

I could go on to enumerate the many fine papers which have recently been published in professional journals,^{23/} but such additional evidence would only serve to emphasize the point which I hope has already been made: there is on the outside of the intelligence community much wisdom and talent which we have neither tapped sufficiently nor used effectively because we are ill organized to do so. This work is relevant to our needs: we cannot afford to duplicate it; we must therefore learn to exploit it in our own behalf.

~~SECRET~~

SECRET

In the past, CIA has had no organization worthy of the name to identify this outside work and to relate it meaningfully to our improvement programs. This gap has been filled with the organization of our Intelligence Sciences Laboratory. This facility is acquiring its own computer and associated equipment to provide an experimental environment closely approximating actual operations. Illustrative of its areas of interest are on-line processing for analysis, pattern recognition, language and text processing, and speech and audio processing. We will thus better bridge the work done outside and our own EDP related operations (see following schematics).

While CIA pioneered much automatic information processing activity with its punched card equipment, our experience with general purpose computer operations is short of six years. In those years we have very considerably expanded our use of these machines.

25X1B

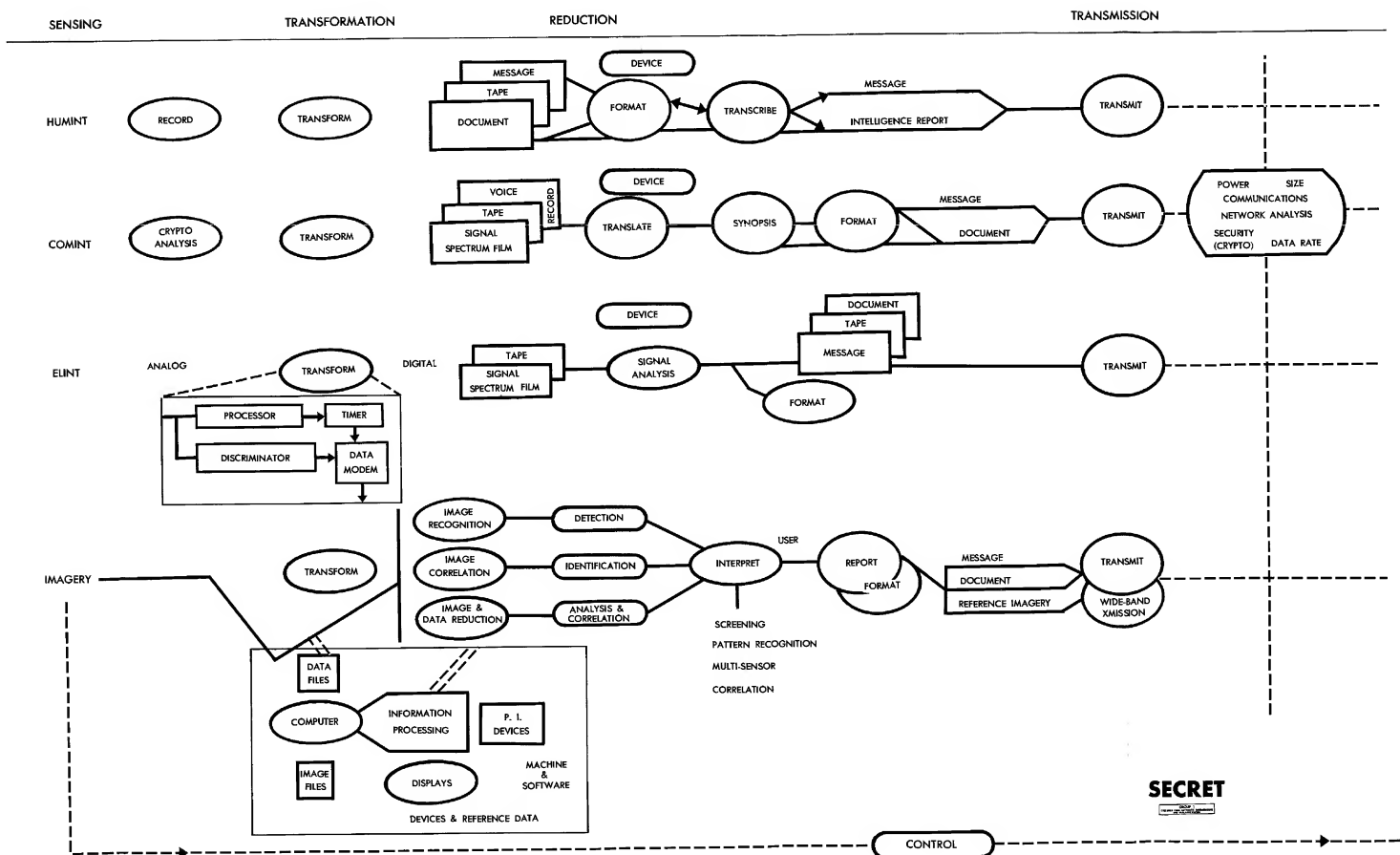
25X1B [REDACTED] The use of the computer in intelligence processing has in the last three years expanded to the point where we see applications in almost every element of the intelligence cycle. Some of the applications and techniques which we consider particularly useful or promising are summarized in APPENDIX B.

A word on how we determine suitable applications may be of interest. There are essentially two basic approaches. In certain large data handling areas, particularly where the information is numeric (military economic costing data, agricultural data, including rainfall and soil moisture statistics) or formatted, computer applications have been initiated by the production analyst with good results. Since data preparation in most cases is done by the EDP staffs, these applications, particularly in the economic area, have produced high-yield products which require little investment on the part of the analyst. A wide range of applications of this type remain to be tried, including computer control of Soviet S&T literature. Projects of this kind, involving primarily collaboration between a production analyst and a representative of the applications division of our central computer facility, can be characterized as special projects.

SECRET

SECRET

AN ILLUSTRATIVE MODEL OF THE INTELLIGENCE PROCESS COLLECTION

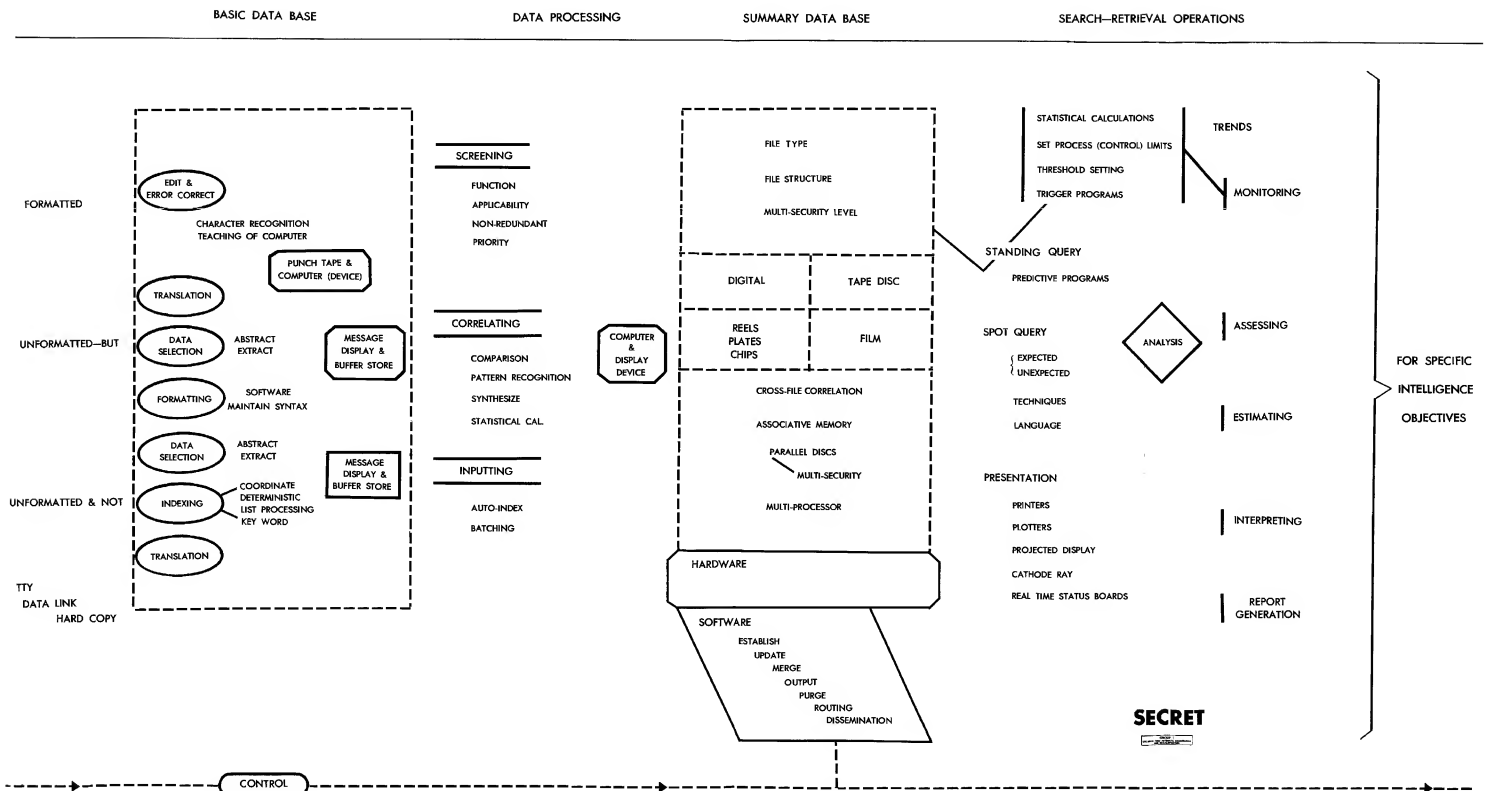


SECRET

SECRET

AN ILLUSTRATIVE MODEL OF THE INTELLIGENCE PROCESS

PROCESSING — ANALYSIS — SYNTHESIS



SECRET

At the other end of the scale is the ambitious attempt to change in a fundamental way the present method of doing business in a major component of the Agency. Our Project CHIVE illustrates this. In one way or another this general project will change the way hundreds of analysts are now working.

There are many unique features of the CHIVE approach. The CHIVE team is a fully integrated group drawn from production, reference, and computer components of the Agency, and includes contract personnel as well. In-house, experienced operators of information systems who have been given training in advanced techniques are in charge. The prospective user of the system is drawn in as active participant. An important geographic area, China, has been selected as the first increment of the system, but will be conducted as a pilot operation in parallel with the old system to permit experimentation before going operational.

The need for CHIVE arose from developments over the past nearly 20 years, during which we evolved a number of special reference services to support the production analyst. The multiplicity of classifications, of indexing tools used for control, of formats employed to collect, disseminate, store and retrieve, made it increasingly more difficult to meet customer needs. The problem of heterogeneity was compounded by the increase in volumes received, and, with the passage of time, the volumes in file. Moreover, the output expected of the production analyst of today calls for greater sophistication within shorter deadlines. Project CHIVE is designed to help him meet that challenge.

Some of the parameters and characteristics of Project CHIVE are summarized in APPENDIX C. Only because computer technology, capability, and capacity are what they are today and will be tomorrow do we dare count upon the success of this project.

And risks there are. This we know. In this costly field it is well to seek ways to reduce risks, but after you have done all you can they are still considerable. We recently had a series of meetings with [REDACTED] during which he shared [REDACTED]

25X1A

25X1A

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

25X1A with us the experience of others in using computers. He reported that of 27,000 systems installed in the U.S. today in some 16,000 installations, 40% are unsuccessful in the use of their computers. This means that 6,500 organizations are not deriving economic benefit or are not achieving their objectives. In 90% of these cases schedules and budgets have been exceeded. The main reason for this in [REDACTED] view is that people using the machines are way behind the technology. In short, machines are much more capable than people are capable of applying them.

My own rule of thumb in the application of machines to non-numeric problems is this: expect half as much in twice the time at twice the cost. If you get it you can count yourself lucky.

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

If we begin with certainties,
we shall end in doubts; but
if we begin with doubts,
and we are patient with
them, we shall end in
certainties.

BACON

V. CHALLENGE OF ASSUMPTIONS

Having searched and studied all possible sources of assistance which techniques, procedures, and machines can offer, the most important job still remains undone, for it is insufficient to improve present means of doing business, or even to find superior means of supplanting present approaches.

What we most need to do is to challenge our assumptions. Reason will prevail and progress be made only when the assumptions by which we live are cultivated, examined, destroyed, and replaced. Assumptions form the skeleton of any information system. They determine its operational role. For that reason we can never ask "why?" too often. We are assisted in this challenge of assumptions by the variety of motives, good and bad, which different members of any organization have. And also by the variety of assumptions held about a given operation. Those who have not viewed an operation favorably will welcome a chance to bring about change. Their operations may be enlarged or improved in some way at the expense of an operation they may view as archaic. Those who have not known much about it may welcome a chance to broaden their knowledge and experience, which in turn may lead to more rapid personal advancement. Those who have a stake in the present operation may welcome a chance to attract greater attention and support to it than it previously enjoyed, particularly should the validity of the challenged assumption be confirmed in the mind of the examiner as a consequence of his close scrutiny of it.

As we all know, in the impersonal catacombs of the bureaucracy, outmoded notions and routines are likely to be jealously guarded. This is human, and indeed management tends to foster such attitudes. We build morale by impressing upon subordinates the importance of diligent discharge of particular tasks performed in particular ways. An abrupt challenge which disturbs these patterns can be as startling as a slap in the face. We need nevertheless to provide for change. Genius lies in doing so in a manner which builds rather than destroys morale.

Here, as in other fields of human endeavor, a sense of participation is the most promising approach. Making the production analyst, for example, an active rather than a passive participant in bringing about change is an educational method of the greatest power. His interest is aroused, he shares in the program's development, he supports its implementation and operation.

The inquiring mind skeptically turned to a situation of long standing can yield commendable and at times even surprising results. We had, for example, a collection program which by 1964 produced 40,000 reports annually. Collecting, typing, editing, processing, recording, and disseminating these so preoccupied case officers that minimal effort went into guiding the collectors.

An analysis of this situation disclosed many useful and interesting things. However, the most significant attack upon this particular citadel was to challenge the assumptions that quantity of reporting was what was wanted; that quantity of reporting best filled information gaps; that quantity of reporting was the way for the post and the collector to get the highest marks; that standard written guidance sufficed as the medium of communication between analyst and collector.

We set quality as the standard, brought collector and user closer together, and relaxed the formalities for reporting management-type information (e.g., statistical data). The disabusing of falsely held notions and anticipations had immediate consequences. In one year's time volume was down 36%. In the second year volume dropped 12% from the new base. Thus in two year's time the number of reports decreased by 44%. This saved time and paper-handling. More importantly it greatly improved quality of reporting.

In another area, an assumption of long-standing has been that given certain facilities designed to support him, the production analyst will in fact use them. And if he doesn't he should be instructed to do so for his own good.

Here we are entering a complex area involving the characteristic inertia in human nature, the futility of legislating human conduct, and relative factors of confidence of an analyst in his own file versus one organized by others in his behalf. Don Swanson^{24/} has suggested the application in such cases of the principle of least action; namely, the design of any future information service should be predicated on the assumption (whether true or untrue) that its customers will exert minimal effort in order to receive its benefits. Furthermore, they will not bother at all if the necessary minimum is higher than some daily low threshold,

In our planning, or in reviewing assumptions, there are of course certain constants which have profound implications. One is the rate at which information can be read by any one person, sometimes referred to as "individual channel capacity." This taken together with the fact that recorded knowledge accumulates through the years means that each individual will see a decreasing fraction of available information. If it is necessary to blanket the accumulation, then the response must be increasing specialization. This leads to "twigging," the phenomenon which occurs in the endless fractionation of interest and knowledge in various fields. The twigs remain the same size but the tree gets larger. In this situation, can we safely assume, as we often have, that adequate coverage of information can be provided by a staff whose size is constant?

Take another situation. Production analysts often request documents five, eight, or even over ten years old. We have assumed that a request of these is an indicator of value, and that we therefore dare not risk purging old material from ever growing holdings.

Just as we have considered it unwise, as has many times been suggested, to use senior production analysts as a filter to thereby reduce appreciably the present flow of information into the system. Should not this assumption too be challenged. Indeed the assumption that lots of paper is in itself bad, and that volume reduction of any kind is in itself good, needs to be under constant scrutiny.

Should we not be more daring in our search for the real edges of our new world? We devote vast treasure to R&D programs which promise gadgetry, accepting willingly the risk of possible failure. Isn't it time we just as willingly seek the promise of experimentation in the softer information processing sciences. And in doing so incur no greater opprobrium for failure in this field than we incur in the development of hardware.

25X1A Recently, at a seminar on "The Computer and the Policy-making Community," [REDACTED] concluded that we had perhaps been taught "...that the vices of computers are our own."^{25/} It is time we learn more about these vices, and, as with intelligence information, get these under better control.

Knowledge is of two kinds.
We know a subject ourselves
or we know where we can
find information about it.

SAMUEL JOHNSON

REFERENCES

1. [REDACTED] The guiding of intelligence collection. IMC Proceedings, November 1958.
2. [REDACTED] On processing intelligence information. IMC Proceedings, November 1958.
3. [REDACTED] Advanced techniques for processing intelligence information. IMC Proceedings, October 1962.
4. [REDACTED] Introductory remarks to the demonstration of machine translation. IMC Proceedings, October 1962.
5. PRICE, DEREK JOHN DE SOLLA. Science since Babylon. New Haven, Yale University Press, 1961. 149p.
6. BUSH, VANNESVAR. As we may think. In the Atlantic Monthly p.101-108, July 1945.
7. THE NEW computerized age. In Saturday Review p.15-37, 23 July 1966.
8. GOOD-BY to Gutenberg. In Newsweek p.85-88, 24 January 1966.
9. Mr. McLuhan's book, UNDERSTANDING MEDIA--The extensions of man, is of interest to intelligence.
10. PLANNING CONFERENCE ON INFORMATION TRANSFER EXPERIMENTS, Massachusetts Institute of Technology, 1965. INTREX: report of a planning conference on information transfer experiments, 3 September 1965. Edited by Carl F. Overhage, and R. Joyce Harman. Cambridge, Massachusetts, MIT Press, 1965. 276p.
11. MILLER, JAMES G. Interuniversity Communications Council, EDUCOM. Unpublished manuscript.

- 25X1A
12. WASHINGTON STATE UNIVERSITY. Proposal for the on-line automation of the Washington State University Library, submitted...to the National Science Foundation, Office of Scientific Information Service. Pullman, Washington, 1966. 26p.
13. [REDACTED] Using computers in a new university [REDACTED] American Library Association Bulletin 59:823-826, October 1965.
14. KING, GILBERT W., and others. Automation and the Library of Congress. Washington, Library of Congress, 1963. 88p.
15. LICKLIDER, J. C. R. Libraries of the future. Cambridge, Massachusetts, MIT Press, 1965. 219p.
- 25X1A
16. BOURNE, CHARLES P. Methods of information handling. New York, J. Wiley, 1963. 241p.
17. [REDACTED] and ROBERT M. HAYES. Information storage and retrieval. New York, J. Wiley, 1963. 448p.
18. GREENBERGER, MARTIN ed. Computers and the world of the future. Cambridge, Massachusetts, MIT Press, 1962. 340p.
19. FEIGENBAUM, EDWARD ALBERT, and JULIAN FELDMAN, eds. Computers and thought. New York, McGraw-Hill, 1963. 535p.
20. BORKO, HAROLD, ed. Computer applications in the behavioral sciences. Englewood Cliffs, New Jersey, Prentice-Hall, 1962. 633p.
21. U.S. NATIONAL SCIENCE FOUNDATION. OFFICE OF SCIENCE INFORMATION SERVICE. Current research and development in scientific documentation. No. 13, November 1964. Washington, 1964. 486p.

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

22. CONFERENCE ON LIBRARIES AND AUTOMATION. Airlie Foundation, 1963. Libraries and automation; proceedings, edited by Barbara Evans Markuson. Washington, Library of Congress, 1964. 268p.
23. Illustrative of the type of papers of which I speak are the following:
- . ASHWORTH, WILFRED. Librarianship and other disciplines. In ASLIB Proceedings 18:152-159, June 1966.
 - . BAR-HILLEL, YEHOSHUA. Is information retrieval approaching a crisis? Washington, Office of Naval Research, 1961. 11p. (ASTIA no.269881)
 - . FIOCK, L. R., jr. Seven deadly sins in EDP. In Harvard Business Review p.88-96, May/June 1962.
 - . FLOOD, MERRILL M. The systems approach to library planning. In Library Quarterly 34:326-338, October 1964.
 - . HEILPRIN, LAURENCE B. On the information problem ahead. In American Documentation 12:6-14, January 1961.
 - . KNOX, WILLIAM T. The new look in information systems. Washington, Office of Science and Technology, Executive Office of the President, 1966. 18p.
 - . OETTINGER, A. G. An essay in information retrieval, or the birth of a myth. In Information and Control 8:64-79, February 1965.
 - . SALTON, GERARD. Progress in automatic information retrieval. Reprint from IEEE Spectrum vol. 2, no.8:90-103, August 1965.
 - . SHAPIRO, E. B. An evaluation of computers in text editing. Menlo Park, California, Stanford Research Institute, 1966. 31p. (SRI project 5849)

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

- . SIMONTON, WESLEY C., ed. Information retrieval today; papers presented at the institute conducted by the Library School and the Center for Continuation Study, 19-22 September 1962. Minneapolis, Center for Continuation Study, University of Minnesota, 1963. 176p.
 - . U.S. FEDERAL COUNCIL FOR SCIENCE AND TECHNOLOGY. COMMITTEE ON SCIENTIFIC AND TECHNICAL INFORMATION. Recommendations for national document handling systems in science and technology. Washington, 1965. 20p. Appendix A--A background study--vol. 1. Washington, 1965. 1v.(various paging)
 - . U.S. NATIONAL ACADEMY OF SCIENCES. NATIONAL RESEARCH COUNCIL. AUTOMATIC LANGUAGE PROCESSING ADVISORY COMMITTEE. Language and machines: computers in translation and linguistics. Washington. In preparation, announced for publication in October 1966.
24. SWANSON, DON R. On improving communications among scientists. In Library Quarterly 36:79-87, April 1966.
25. SCHWARTZ, JUDAH L. Computers and the policymaking community. Livermore, California, University of California, Lawrence Radiation Laboratory, 1966. 19p. (AEC contract no.W-7405-eng-48) (UCRL-14887; TID-4500, UC-32)

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

Nothing can come out of nothing,
nothing can go back to nothing.

PERSIUS

SELECTED COMPUTER APPLICATIONS AND TECHNIQUES

Collection

The computer is now being used in a rudimentary fashion in collection management and in determining collection system performance. We see, for example, the "bookkeeping" functions of a requirement registry performed by the machine. More sophisticated applications are on the horizon which involve the testing of hypotheses which a manager of intelligence collection normally is concerned with, i.e., given certain collection targets and priorities and specific collection sources, how might these collection resources be allocated against these targets? At this point, however, the computer is not used as an integral part of this optimizing process; rather, it provides an answer to a specific computational translation of a collection hypothesis.

Data Reduction

A broad view here requires a look at areas which we consider normally under data reduction--signal analysis and the like--as well as other activities which not only filter data, but present them in a form that a "lay" analyst can use effectively. The computer has played an important role in the former area for quite some time. We see, however, demands for more sophisticated processing of higher volumes of raw data--better precision in wave form analysis, ability to perform more sophisticated spectral analysis and correlation, a wider variety of emitters to identify, and so on. In the latter area, we should consider topics such as automatic language translation, automatic filtering of redundant textual material, automatic dissemination, and so on. We see little operational impact of computers in these areas aside from long-term experiments with limited chance of payoff.

Retrieval and Analysis

Again, we can cite two subcategories: computational and non-computational data processing; and, as before, progress in the former is considerably advanced over the latter (this is independent of any value judgments on the utility of the product involved). For example, the

computer has been used very effectively in filling in the complete picture of a missile or space event with only fragmentary data available. Secondly, where a particular activity of intelligence interest can be reduced to a mathematical model, some successes have been achieved--one example here is our ability to analyze a particular air defense system.

In the area of non-computational data processing, the computer has been used principally as a file updating, sorting, searching, and printing device. Examples here include files on aircraft movements, personnel travel, facility status, and so on.

Some success has been achieved in assisting economic intelligence analysis, which is a combination of both types of data processing. Some of the first Agency applications of computers to intelligence were in this area and continue to be of some importance--computation of Soviet military expenditures and of Soviet gross national product.

Production

The process of providing intelligence material in a form suitable for wide dissemination involves considerable expense and in some areas is amenable to computer assistance. The two particular examples are cited here. The production of high-quality, printed material is now being assisted through the use of a computer-generated tape which is fed to an automatic photo-composing device. New editions of the National Intelligence Survey will be printed using this technique without sacrificing the aesthetic qualities of this publication (the use of multiple fonts and other complexities in print composition). Secondly, the computer is now being used on an experimental basis to provide assistance to cartographers. It is intended that a large bank of cartographic information be machine stored and selectively retrieved and plotted to provide the basic map over which specific intelligence information would be superimposed. Considerable flexibility would be available in terms of scale, projection, degree of detail, and number of geographic features shown on the map.

A basic criterion that might be applied to determine if the production analyst is making effective use of the computer is, "Would there be any disastrous effects if the computer were suddenly withheld from any of these application areas?" A candid answer in most cases, unfortunately, would be no. A basic reason for this is that the analyst has maintained separate backup facilities in many cases--perhaps because of his reluctance to rely so heavily on the computer. Secondly, several of these applications are in the "why don't we try using a computer" category--without appropriate analysis of the costs, effort, and ultimate payoffs involved. There is some cause for optimism, however. Notable successes have been achieved in cases where (1) the analyst initiates the requirement, (2) he understands the capabilities as well as the limitations of the computer, and (3) he permits the computer specialist to get deeply involved in the early stages of problem analysis.

Input

Considerable attention is now being given to the concept of "source data automation." This implies the keying of data as close to the source as possible. In some circumstances it implies direct entry of these data into the machine. A technique of interest here is optical character recognition, where the emphasis has recently shifted from an interest in reading text to the use of character readers in the place of punching paper tape or cards. Where reasonable quality control can be applied to the typewriter font, paper contrast, and character registration, available print readers are now competitive as an input device with punch card or punch paper tape readers. This input preparation technique has several obvious advantages; the use of standard office equipment for keying, decentralization of input preparation, the ability of the human to read exactly what the machine reads, ability to implement a "turn-around document" concept, i.e., computer printouts can be modified and then fed directly back as input.

Progress is also being made in design of economical secure tape typewriters. In the next few years, we expect to see such devices available for under \$5,000. It is also reasonable to expect that the cathode ray tube or

typewriter terminal will be used as a direct data entry device for large volume inputs. This would provide a facility for direct feedback of detected error conditions which could be serviced on the spot.

Processing Units

The basic design of computers has remained unchanged since the invention of the stored program concept over 15 years ago. It was designed for arithmetic operations and as such is not particularly appropriate to the solution of non-numeric data processing problems. However, attempts to design a machine which better fits these problems have not been very successful. For example, parallel search memories where all cells are interrogated simultaneously have not progressed beyond the laboratory. Attempts have also been made to build a machine which processes materials sorted in something called an "information processing language" (IPL), where data elements are stored in a tree structure and search progresses up and down branches of the structure. Another example of a special-purpose machine is one currently installed in the CIA Computer Center called the Automatic Language Processor, which employs a large-capacity disk called a "photostore" which is used as a dictionary for lexical processing. Simulations of all of these special machines have indicated that the general-purpose device which is now available is an effective and relatively inexpensive--if unsophisticated--facility for the solution of data processing problems. The principal reason for this is that the market for the general-purpose computer is extremely large and competitive, and the cost for a given amount of processing continues to decrease.

Files and Storage

The comments made above are applicable here as well. That is, storage devices normally made available with general-purpose machines cannot be considered optimal for most intelligence data processing problems. But the capacity and speed of these devices makes them economically feasible. While direct access devices such as storage drums and disks have been available for several years, our ability to use them effectively is still behind the

electromechanical technology. For example, we find that the various segments of disk storage are really considered to be small magnetic tapes--data within each segment are usually searched serially.

Software

This term was coined about four years ago to identify a concept which implies that many of the facilities that must be available to users of computing equipment are appropriately implemented through a set of programs rather than through specific hardware logical features. For example, while a computer contains a control unit which automatically determines the sequence of its operations, more sophisticated automatic control features, such as automatic scheduling of major jobs on the computer, have been assigned to programmers for implementation. Similarly, more sophisticated languages for commanding the machine--other than the basic instruction set which includes functions such as add, subtract, move, and so on--have obvious advantages to the user. These again have been implemented through programs which translate these command languages into the elemental instruction set available directly to the hardware.

This concept has mushroomed to the point where buyers of computer systems demand a wide variety of control and utility programs as an integral part of the package to be rented or bought from the manufacturer. The software concept, along with the increased capacity and speed of computing devices, has resulted in the notion of the computer as an information processing utility. As a result, we hear the term "time-sharing" quite often when we talk to computer people these days. In this kind of an environment, the computer user sitting at the console has the illusion of the complete computer being made available to him continuously. In reality, however, he is only given short bursts of the time available on a computer. These time slices might be several milliseconds for each of perhaps 100 users. This area is similar to the method of interleaving telephone conversations in a modern communication system. If this concept is workable--and experiments thus far show promise--the computer user, whether he be a programmer or intelligence analyst, will have more direct access to computer facilities.

Analog Devices

The digital computer has for quite some time surpassed the analog computer in its speed and precision. However, an increasing amount of data which must be processed in an intelligence environment is collected in analog form. If the digital computer is to be used for the analysis of these raw data, more efficient means of converting the analog signal to digital form must be available. The principal limitations of conventional converters are speed, and quantization precision. In addition, it would be desirable if the analog portion of such a converter would have feedback facilities for automatic retuning of the device to enhance the quality of the signal being processed. Converters which incorporate some of these advantages and overcome some of the speed and precision limitations previously encountered have been designed and will be used soon experimentally. The results of these experiments should be followed closely because of the vital role that analog collection systems promise to play in intelligence in the future.

In the library of the future, man
will continue to read books, gain
insights, think, and make discoveries.
But the library will do most of the
searching, transforming, interpreting,
and checking of information he needs,
and thereby will free him for more
creative uses of stored information.

BOLT BERANEK & NEWMAN

SECRET

CIA'S PROJECT CHIVE

CIA's information processing problems are being addressed by a combined team consisting of representatives from the Offices of Central Reference and Computer Services on a project basis (Project CHIVE).

This project is a two-pronged effort to achieve an improved capability to store and retrieve intelligence, through the combination of machine and human techniques for managing and controlling massive volumes of information. First, a new concept of Agency-wide information service is being developed with the emphasis placed on a fast integrated response to analyst queries using a reasonable balance between the intellectual capabilities of intelligence analysts and the data storage, selection, correlation, and display capabilities of electronic equipment. The data base for the all-source system will be comprehensive, concentrating on named object or "hard" intelligence facts which will be used for document retrieval, information file building, and (eventually) for automatic correlation--or the retrieval of data through inferential logic.

From the user point-of-view, the CHIVE system will have unique characteristics which should provide a number of advantages. Basic features of the system are:


- . All-Source Retrieval--Establishment of an integrated, all-source file system covering every type of printed document of use to the Agency, including maps and photos, at whatever classification level. Here, the significance to the user would be that he would be able to get a total system response, assured that all available sources have been tapped and their relationships have been exploited.

- . Single-Point Service--Reorganization of the Agency's central information retrieval service under a radically different management umbrella--that is, by geographic area--with each geographic element exploiting all available source material and using common indexing and procedural standards. As a result, the user would need to contact only a few people to bring the full power of the system to bear on his problem.

SECRET

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

- . Literature Searching--Fulfillment of the basic function of serving as the Agency's institutional memory by responding to such requests as: "Give me all books, documents, reports, etc., that concern Soviet anti-ballistic missile research." The response in this case would be relevant documents, not the answers themselves.
- . Retrieval of Basic Facts--Answers to specific questions derived from the accumulation of data stored. The latter will include a host of facts concerning foreign personalities of all kinds, organizations (political, cultural, scientific, educational, etc.), installations and other fixed facilities, and activities and events (for example, trade, subversion, missile firings, etc.). Thus, the user will be able to ask, and have answered inside of 15 minutes, such questions as:
 - "Who is the head of the Soviet Rocket Forces?"
 - "When did Mao Tse Tung last appear in public?"
 - "What operational ICBM sites have located in the Baku area?"
- . Counts, Correlations, and Trend Analysis--Answers to questions, both for management control as well as substantive intelligence purposes, such as:
 - 
 - "How many papers on solid fuel rockets were published by the Chinese in 1965?"
 - "How many CIA intelligence reports were issued on Haiti in July 1966?"

25X1B

Given the ability to make counts, we can also observe how such counts change as a function of time and thus estimate trends. For example, we might observe that there is a marked deviation from the norm in the number of appearances of Chinese leaders, and so forth.

SECRET

SECRET

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

- . Detection of Redundancies and Inconsistencies-- Identification of duplicative information in the system store as well as apparent contradictions which might suggest follow-on research to be undertaken by intelligence analysts.
- . Automatic Inference-Making-- Assuming we can devise a system that can store a wide variety of facts about people, institutions, and activities, and can then "operate" on these data, it should be possible to infer information that man x has worked on thermonuclear fusion, and if man x is currently living in city y, we might infer a certain probability that thermonuclear research is being conducted in city y. The variety of problems to which such a man-machine capability might be applied is largely a function of the imagination and ingenuity of the intelligence analyst.
- . Machine-Assisted Language Translation-- More comprehensive and rapid exploitation of the totality of information received in foreign language documents through (a) actual machine translation of Russian language documents and (b) a machine system for rapidly converting oral translations of other languages, via the Stenographic process, to printed documents in English.
- . Analyst Referral Service-- Establishment of a central directory or "profile" of human and other information resources scattered about the Agency possessing expert knowledge in specialized areas. Such a tool would be used to supplement the CHIVE system by identifying analysts as well as intelligence problems.
- . Micro-Image Document Store-- Installation of a system for storing the vast quantity of Agency documentary receipts (up to 1.5 million documents/yr.) in a miniaturized form thus reducing space requirements and providing a random-access, rapid retrieval and reproduction capability. While initially the storage medium will be aperture cards, it is planned that the system will be converted by 1970 to a large-capacity, document storage device on the order of a second-generation "Walnut."

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

SECRET

- . Remote Querying--Provision of remote communication capabilities, coupled with a flexible file structure and machine command language, which will enable users to interrogate and maintain data files directly and on an up-to-date basis. Special-purpose files maintained by analysts could eventually be included in the central system. Ultimately, we would hope the system would lighten the individual analyst's burden in file housekeeping and serve as a more effective extension of his memory.

We see the development of an improved information storage and retrieval capability as extending over two five-year phases. During the first five years, we will concentrate on implementing the full CHIVE system in house through a planned evolution by geographic area, beginning with Communist China. During the next five years, the goal will be to consolidate the CHIVE system with other specialized Agency information systems with the idea of achieving an integrated Agency information network. In parallel during the second phase, we expect to achieve integration of Inter-Agency EDP information systems and to provide a world-wide communications network in support of U.S. intelligence activities and operations.

Illustrations

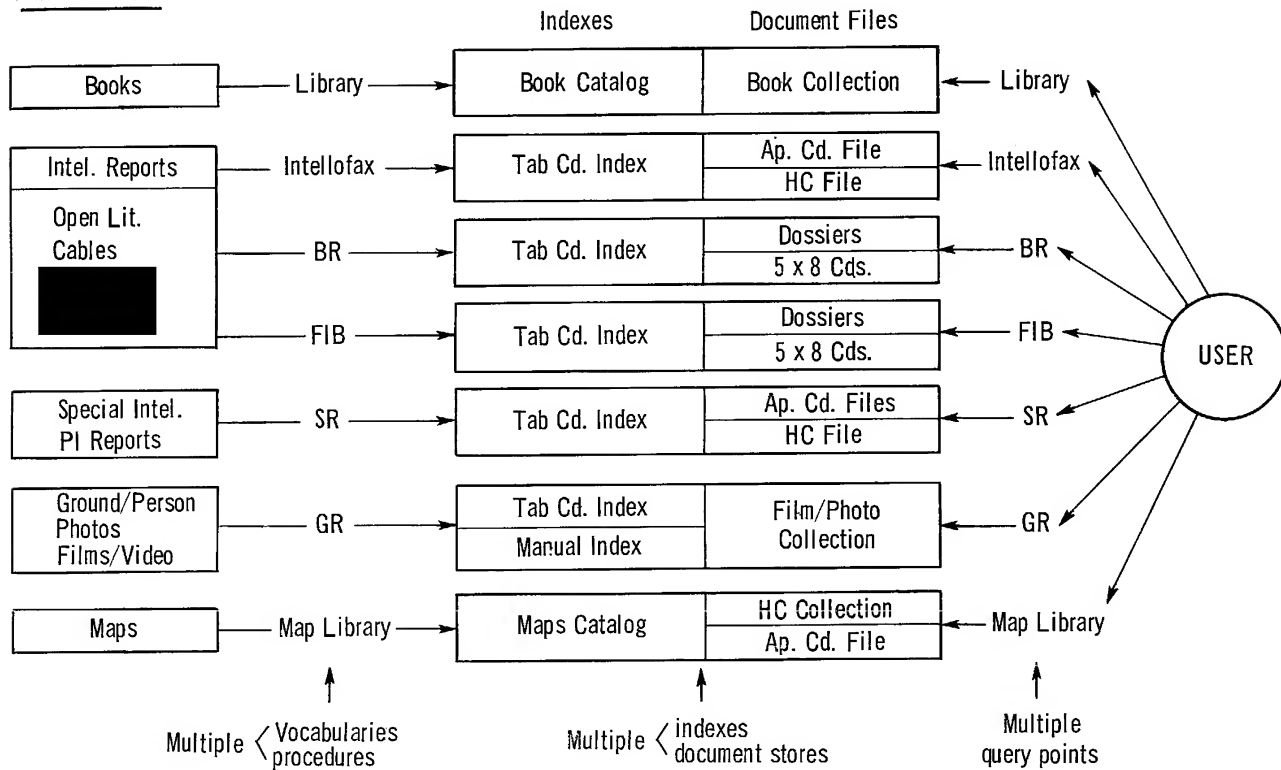
- . Current System
- . CHIVE System
- . CHIVE Functional Overview
- . CHIVE Information Language

CURRENT SYSTEM

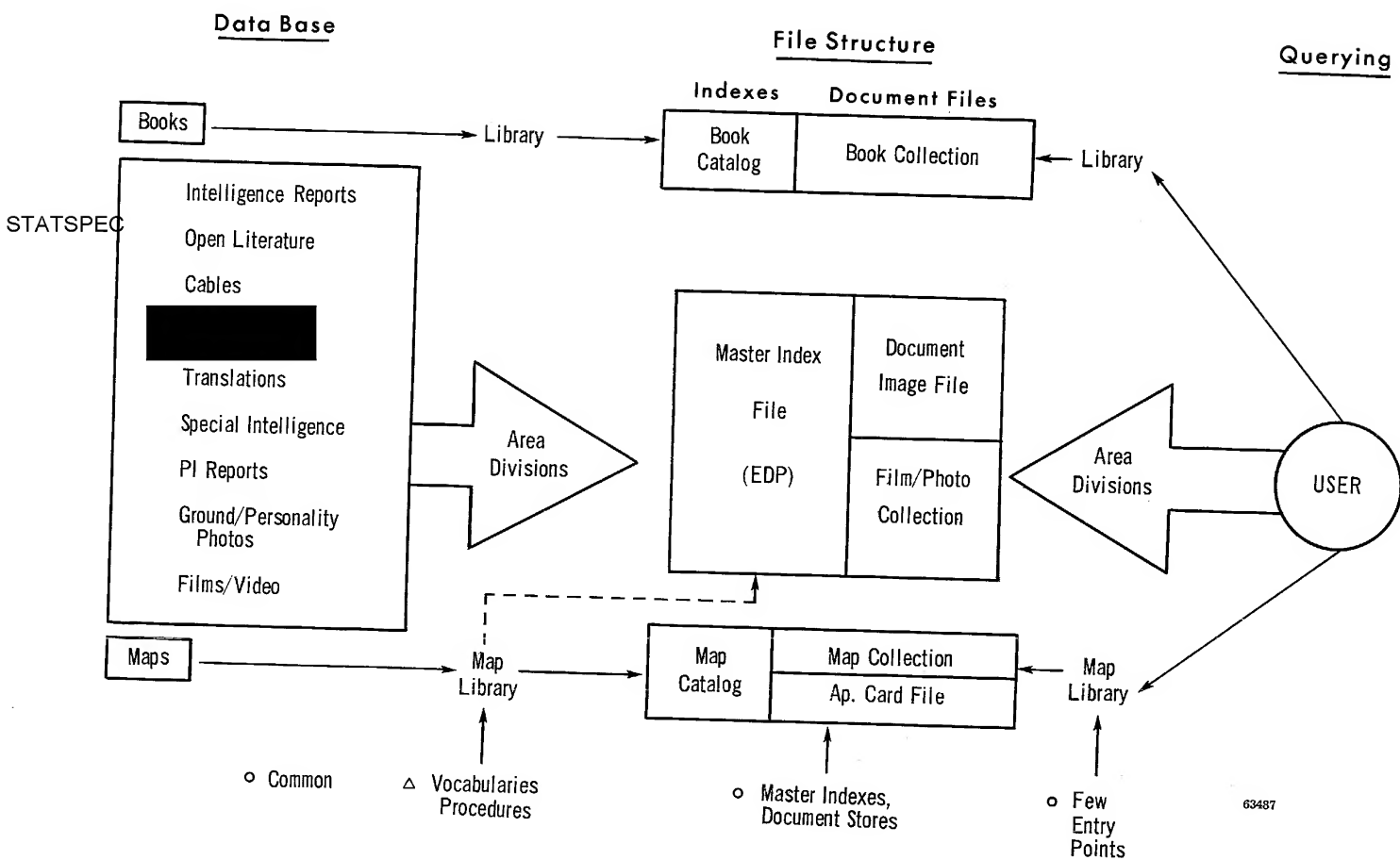
Data Base Documents

File Structure

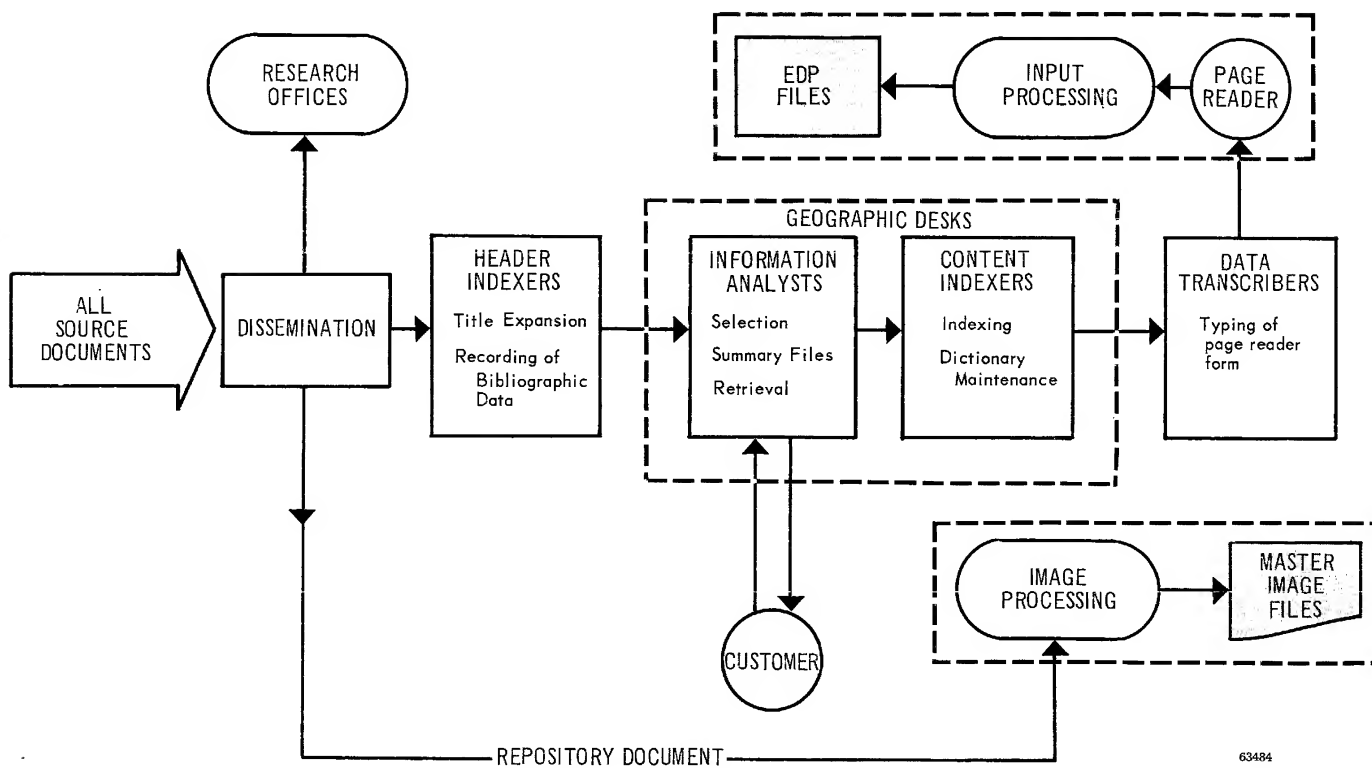
Querying



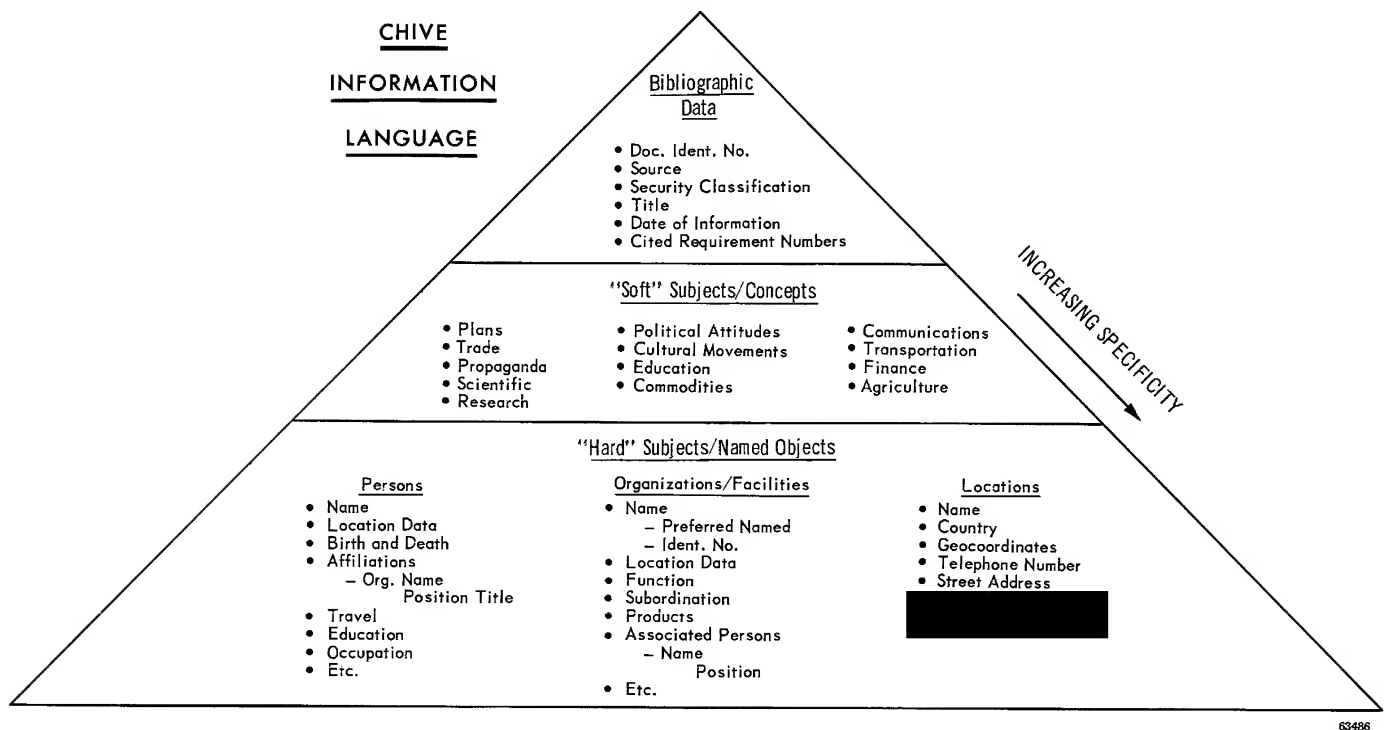
CHIVE SYSTEM



CHIVE FUNCTIONAL OVERVIEW



63484



Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6

Approved For Release 2001/11/22 : CIA-RDP80B01139A000300110003-6